Patent Idea Details for Idea #41827

General Information

Title: Method for Achieving Scalable Control H.323 Signaling
ID: 41827
Patent No.: ---, ---
URL: [Application No. ---]
Inventors: Leo Nieuwesteeg (leo) and Randall Baird (rbaird)
More details on these inventors listed below.
Date Entered: [redacted]
Date Modified: [redacted]
Date Filed: ---
Date Issued: ---

Background: The ITU's H.323 standard provides a mechanism for establishing voice, video, and other multimedia calls over an IP internet or other packet medium. The H.323 signaling (that is, the information which controls the call throughout its life) is handled by two different exchanges of PDUs, both mounted on top of TCP. In H.225.0, a set of PDUs modeled after Q.931 is exchanged to provide basic call control. In H.245, PDUs are exchanged to control media, codec, and conference selection, and to make modifications to these selections while a call or conference is ongoing.

In H.323 v1, both the Q.931 and H.245 connections require two separate TCP connections, which must persist during the life of the call. In H.323 v2, Q.931 and H.245 can in some cases be multiplexed onto a single TCP connection.

A common mechanism for allowing third-party call control in H.323 is to route the Q.931/H.245 connections through an intermediate call agent, using either gatekeeper-routed call signaling or an H.323 proxy. In the GK-routed call signaling mechanism, an endpoint requests to be connected to an address via a directory and admission module called a gatekeeper (GK). The GK, rather than instructing the endpoint to connect directly to the terminating endpoint, instructs that the connection be made through the GK itself. The GK is then capable of controling and modifying the signaling without affecting the the underlying media streams, which can flow directly between the endpoints. If a proxy is used, the GK can instruct the originator to connect the Q.931 and H.245 to the proxy, along with the underlying media stream. Proxies are typically used when the media must be converted for firewall or NAT traversal, or when the QoS of the media must be modified.

Note that H.323 explicitly allows for differing paths for the signaling and the media streams (which are transported using RTP). This allows signaling control elements to be inserted into the path without adversely impacting media performance. Call agents must obviously have a mechaism through which endpoints like gateways can be instructed to emit appropriate RTP streams to other endpoints. This can be accomplished using a protocol which is much simpler and more efficient than H.323, like SGCP or MGCP.

The problem arises when it is necessary to have massive numbers of H.323 calls flowing through a single call agent. Since each call can consume up to 2 TCP connections and most platforms have limits on TCP connections that are well below 10,000, only a maximum of less the 5000 H.323 calls can be controlled by a single call agent box. In a carrier-class implementation based on H.323, hundreds of thousands or even millions of calls must be controllable at any given time.

Summary: Scalability can be achieved by overcoming the TCP limitations. To do this,

[redacted] actually perfo[redacted]ing their own control functions, they backhaul the signaling to a centralized call agent. The signaling backhaul has two major properties.

First, the backhaul mechanism multiplexes the signaling for all of the Q.931/H.245 sessions onto a single session between the backhaul proxy and the call agent. This reduces the need for thousands of TCP connections on a call agent down to merely needing a single session between the call agent and each backhaul proxy.

Second the backhaul protocol need not be TCP. Using a different type of transport, dropped TCP packets pertaining to any subset of the multiplexed connections need not block the transmission of PDUs for the rest of the multiplexed set of connections. Several transports can achieve this capability. (See Cisco Use section for how we do it.)

When the call agent receives the mutliplexed PDUs for Q.931/H.245, they can be demuultiplexed, intepreted, and acted upon as if they were received directly over TCP. If other PDUs must be sent to a peer H.323 entity in response to the PDUs coming from the backhaul proxies, the call agent "front-hauls" them by sending them over a similar multiplexed session back to the backhaul proxies.

When the backhual proxy receives front-hauled traffic from the call agent, it demultiplexes it and forwards it to the peer H.323 endpoint over normal TCP connections. In this way, all of the behavior of H.323 can be reproduced without the scalability problems.

Advantages: 1) Scalability is improved. While it is certainly possible to distribute the call agent intelligence across multiple platforms, many carrier-class solutions require extreme fault tolerance with very expensive hardware. Backhauling the H.323 traffic to a single, highly fault tolerant node can have significant cost advantages.

2) If H.323 is running on embedded gateways and other terminals, the range of functionality may be quite limited. Backhauling the traffic to a central, highly flexible call agent allows features to be added easily and rapidly.

Cisco Use: The Nomad project is implementing a backhaul system using the RUDP and RLM protocols. Nomad is currently a 2600 which is capable of backhauling SS7 and NI-2 signaling stacks. It could be easily extended to backhaul H.323 signaling.

The Virtual Switch Controller currently acts as a call agent for SS7, NI-2, and other protocols. It could also be easily extended with H.323.

There are two major applications which could benefit from the VSC/Nomand combination handling H.323. In the first, the VSC would control a large bank of MGCP-controlled voice gateways and would communicate with a large bank of Nomads to generate H.323 signaling. The H.323 sessions are usually necessary for inter-carrier applications. For example, when carrier A needs to terminate a call in carrier B's network, H.323 is a natural mechanism, since the standard eases complex interoperability, both inter-carrier and inter-vendor. (Note that for intra-carrier transit networks, it usually makes more sense to use MGCP to control both the ingress and egress gateways, which then can spew raw RTP streams at each other without incurring the cost and complexity of using H.323.)

Another, similar, application occurs when a carrier requires that the structure of its POPs be hidden from other carriers, or when inter-carrier QoS requirements vary. In this case, it is desirable to do a full media proxy at the carrier's border. In this configuration, the VSC uses MGCP or SGCP to control the flow of RTP packets between the voice gateway and the border proxies. Only when the call is routed to the other carrier does the VSC use a bank of Nomads to backhaul the necessary H.323 signaling so that the inter-carrier stream appears to be a fully compliant H.323 session.

[redacted]